It is not easy to establish and reason a new MDDS that is adequate of the needs of the medium and small clients but it is worth to invest some efforts in it. Taking into account the discussion above, we could define the following features of such model:

- It has to guarantee the **independence of the user** from the technologies, i.e. the user have to be free to chose for each component of the solution the existing technology that is the best for this component;
- It has to guarantee that the user will obtain a service with a **quality that is relevant to the paid cost**;
- It will be very good if the model has the **tolerance for the qualification** of the user and to allow, etc.

In order to prove feasibility of such model a lot of programming and experimental work should be done. Each feature of the model has to be examined on sample applications. For example, implementing of efficient and fast interoperability of technologically non-homogenous components is a great challenge.

Bibliography

[CIO-11,2006] Editorial, CIO, No. 11, 2006, Internet version. http://www.cio.bg/?call=USE~home;&page=paper&n=1148.

[CIO-3,2007] Editorial, CIO, No. 3, 2007, Internet version. http://www.cio.bg/?call=USE~home;&page=paper&n=1391&pn=2

[O'Reilly,2004] http://tim.oreilly.com/articles/paradigmshift_0504.html

Authors' Information

Neli Maneva – Senior Research Fellow; Institute of Mathematics and Informatics, BAS, Acad. G.Bonthev St., bl.8, Sofia-1113, Bulgaria; e-mail: neman@gbg.bg

Krassimir Manev – Associated Professor, Faulty of Mathematics and Informatics, Sofia University, 5 J. Bourchier str, Sofia-1164, Bulgaria; e-mail: manev@fmi.uni-sofia.bg

THE LOCALIZATION PROCEDURES OF THE VECTOR OF WEIGHTING COEFFICIENTS FOR FUZZY MODELS OF CHOICE

Elena Prisyazhnyuk

Abstract: The author analyzes the localization procedures of the vector of weighting coefficients which are based on presenting the function of value by additive reduction adapted to fuzzy models of choice

Keywords: fuzzy sets, coefficients of value.

Introduction

Different kinds of uncertainty are to some extent characteristics of practically any situation of making decision in which the expert information is used. The result of the research [Ларичев, 2002; Борисов, 1989] show that the main difficulty is caused by the necessity to appraise numerical values of the objects (variants, criteria) or to give numerical evaluation on the ratio scale between them. It is known that verbal definitions usage allow to define the preferences of a person who makes decisions (PMD) more steadily. This approach seemed to be even more justified because in prevailing number of the cases it is enough to have the approximated characteristic of the data set and the expert info usage does not demand high accuracy.

The experts' evaluation in fuzzy models of choice is described by the functions of belonging to the fuzzy set. The functions of belonging can be interpreted differently: as "subjective probability", expert's confidence degree in object's belonging to the concept described by fuzzy set, the opportunity of its interpretation by this concept and so on [Борисов, 1989]. The choice is characterized by preference relation R the meaning of which in fuzzy models consists of the fact that it can point out for every two objects:

- the fact of preference of object α[1] over α[2]. The preference function μ_R(α[1], α[2]) in this case is substantially interpreted as the expert's degree of confidence in the fact that α[1] is not less preferable than α[2]. The confidence degree can be described both numerically and verbally, for example, by the linguistic variables

"confidence degree"={very low, low, average, high, very high}. The linguistic preference relation of this type corresponds to the decision making situations when PDM doubts in preferences existence in relations of the certain objects and because of that it's difficult for him to express them only in terms "yes" (definitely dominates) or "no" (is definitely dominated);

- preference power of object α[1] over α[2]. Fuzzy relation R in this case shows the idea of preference power; besides PMD is sure in the very fact of preference (and in this sense his preferences are clear). Here μ_R(α[1], α[2]) can be interpreted as degree with which α[1] is definitely better (preferable) than α[2].

The problem of the objects' evaluation in terms of value theory settles down to the problem of axiomatic argument and its value function creation. The classical methods used for defining value which represents binary preference relation R($U(a^1) \geq U(a^2) \Leftrightarrow a^1 R a^2$ для $\forall a^1, a^2 \in A$), are generally pretty hard. The basis for their usage, in particular, is a sufficient condition of its existence which is set, for example, by Debre theorem [Пономаренко, 1994]: the preference relation must be complete, reflexive, transitive and continuous, the set of decisions – connected. If Debre theorem conditions are not completed and the function of value which introduces relation R does not exist, it's difficult to use the classical methods.

The procedure of the problem formalization with the help of exchange of fuzzy "vector value evaluation" by additive reduction is suggested.

The task set

Let's A as the universal (distinctly described) set of objects α[j], $j \in J$, where *J* – set of indexes of objects. Each of objects $a^j \in A$, $j \in J$ is characterized by the set of parameters $a^j = (a^j{}_1, ..., a^j{}_i, ..., a^j{}_n)$. Let's mark the set of indexes of objects' parameters *I*, $I = \{1, ..., n\}$. Each object, $a^j \in A$, $j \in J$ corresponds to its vector evaluation in the dimension of objects' parameters Ω^n.

Further on we'll analyze not the set of the values of objects' parameters $\alpha^j{}_i \in$A, $j \in J$ itself but the corresponding set $\omega(a^j{}_i)$, $i \in I$, $j \in J$, where ω – some monotonous reorganization which defines the degree of quantitative characteristics declination from the optimum meanings for each parameter $a^j{}_i$, $i \in I$, $j \in J$ and reorganizes all the meanings of objects' parameters towards the normal type in interval [0,1].

Let the expert consequently define his preferences on the set A as fuzzy binary relations of preference R.

The following approach to the task solution is suggested: we suppose that evaluating the object the expert (consciously or subconsciously) means its vector value. If we consider "vector" function of value as fuzzy additive reduction the task is considered as defining weighting coefficients of reduction (1) – (2):

$$\sum_{i \in I} \rho_i \omega(a_i^1) < \sum_{i \in I} \rho_i \omega(a_i^2), \quad (1)$$

$$\rho = (\rho_1, ..., \rho_n), \ i \in I, \ \rho_i > 0, \ \sum_{i \in I} \rho_i = 1, \quad (2)$$

where (2) – normal vector of relative importance of objects' parameters for the expert's statement about the fuzzy preference relation between the objects.

So, the task is in "localization" (defining intervals of changes) of weighting coefficients of additive reduction (1)-(2). Similar task was analyzed in [Волошин, 2003], in this article we suggest the generalization of the method for fuzzy models of choice.

The localization procedures of the vector of weighting coefficients

Let the person who makes decisions (PMD) considers the object α[1] more preferable than object α[2], and μ – preference degree, $\mu \in [0,1]$. We'll define it by $\mu_{\succ}\left(a^1, a^2\right)$. Then $1 - \mu_{\succ}\left(a^1, a^2\right)$ – preference degree of α[2] over object α[1]. Let's use heuristics: we think that righteousness of inequality (1) comes from PMD subjective idea about object preference α [1] over α [2]. Then for the case of fuzzy preference relation $\mu_{\succ}\left(a^1, a^2\right)$ between objects α[1] and α [2] we'll consider right:

$$\mu_{\succ}\left(a^1, a^2\right) \Leftrightarrow \sum_{i \in I} \rho_i \frac{\omega(a^1{}_i)}{\mu} \leq \sum_{i \in I} \rho_i \frac{\omega(a^2{}_i)}{1-\mu} . \tag{3}$$

It is necessary to create on the basis of preference relations on the set of effective objects A which are consequently defined by PMD, the intervals of allowed values of objects' weighting coefficients (hyper parallelepiped of weighting coefficients - HWC) as:

$$\rho \in K = \prod_{i \in I} [\rho_i^H, \rho_i^B], \quad \rho = (\rho_i,\ i \in I), \quad 0 < \rho_i^H \leq \rho_i^B < 1, \tag{4}$$

$$\sum_{i \in I} \rho_i = 1, \rho_i > 0, i \in I . \tag{5}$$

It is supposed that defining preferences PMD consequently, particularly, in his ideas sets the preference relations between the objects. They satisfy the qualities of transitiveness. Binary comparison is given with the help of the line reduction (1) and HWC as the result of the suggested procedure functioning is reduced step by step ($K^{S+1} \subseteq K^S$, $s = 1,2,\ldots$). So transitiveness of considered binary relation is preserved.

For reorganization of all the meanings of objects' parameters a^j_i, $i \in I$, $j \in J$ to the unlimited kind in the interval [0,1] such formula is suggested:

$$\omega(a^j{}_i) = \frac{a^{opt}{}_i - a^j{}_i}{a^{opt}{}_i - a^0{}_i}, \tag{6}$$

where α $^j{}_i$∈A, $i \in I$, $j \in J$; α $^{opt}{}_i$∈A, $i \in I$ - the best meaning of i-parameter on the set of effective objects; α$^0{}_i$∈A, $i \in I$ - the least meaning of i-parameter on the set of effective objects. Let's consider that α ^{opt} and α ^{0} can be set directly by PMD or defined as maximum (minimum) parameters values which are achieved on the set of admitted.

Taking into consideration (6), the generalized criteria which reflects the total declination of j-object, $j \in J$, from optimum meanings will be presented as

$$D(a^j, a^{opt}) = \sum_{i \in I} \rho_i \omega(a^j{}_i) = \sum_{i \in I} \rho_i \frac{a^{opt}{}_i - a^j{}_i}{a^{opt}{}_i - a^0{}_i}, \quad j \in J .$$

The last formula is the proximity metrics of values vector closeness of object's parameters α ^{j}∈A, $j \in J$, to some ideal (optimum) values vector α ^{opt} = (α $^{opt}{}_1$, α $^{opt}{}_2$,… α $^{opt}{}_n$), weighted in the dimension of parameters. Formula (3) for fuzzy preference relations between objects $\mu_{\succ}\left(a^1, a^2\right)$ will be presented as:

$$\frac{D(a^1, a^{opt})}{D(a^2 \ , a^{opt})} = \frac{\sum_{i \in I} \rho_i \omega(a^1{}_i)}{\sum_{i \in I} \rho_i \omega(a^2{}_i)} \leq \frac{\mu}{1-\mu} .$$

Last inequality can be interpreted in the following way: the statement "object α[1] is preferable than object α [2] with preference degree μ" means that in the dimension of objects' parameters the point which corresponds to object α[1] is located closer to the ideal point than the point which corresponds to object α [2] with degree $\frac{\mu}{1-\mu}$.

Statement 1. Objects $a^1 \in A$ and $a^2 \in A$ are called equal if in "equable" dimension of parameters Ω^n the corresponding points are located within equal distance from the point which corresponds to the ideal object.

Statement 2. Object $a^2 \in A$ is called μ-equal to object $a^1 \in A$ if in "equable" dimension of parameters Ω^n points $\omega\left(a^2\right)\frac{\mu}{1-\mu} = \left\{\omega(a^2{}_1)\frac{\mu}{1-\mu}, \omega(a^2{}_2)\frac{\mu}{1-\mu}, \ldots, \omega(a^2{}_n)\frac{\mu}{1-\mu}\right\}$ and ω(a^1) are equal.

Statement 3. Weighting coefficients of parameters $\rho = (\rho_i,\ i \in I)$ which correspond to μ-equal objects in the dimension of preferences R^n define the intervals' limits of weighting coefficients of objects' parameters.

Argumentation. Let's define sets of indexes of objects' parameters $a^1 \in A$ and $a^2 \in A$ through $I_1 = (i : \omega(a^1{}_i) > \omega(a^2{}_i)) \neq 0$, $I_2 = (i : \omega(a^1{}_i) \leq \omega(a^2{}_i)) \neq 0$, $i \in I = I_1 \cup I_2$. We can renew inequality (2) taking into account defining sets of indexes in the following way:

$$\sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^1) + \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^1) \leq \frac{\mu}{1-\mu} \sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^2) + \frac{\mu}{1-\mu} \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^2)\,. \tag{7}$$

Then the condition of μ-equality of objects α^1 and α^2 will be given as

$$\sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^1) + \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^1) = \frac{\mu}{1-\mu} \sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^2) + \frac{\mu}{1-\mu} \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^s \omega(a_i^2)\,. \tag{8}$$

We can pass over from (7) to (8) if we increase weighting coefficients of parameters which belong to the set of I_1. At the same time we reduce weighting coefficients of parameters which belong to the set of indexes I_2. So weighting coefficients of parameters ρ_i, $i \in I_1$ achieve their upper borders and weighting coefficients of parameters ρ_i, $i \in I_2$ achieve correspondingly their lower borders. As $\omega(\alpha^1{}_i)$, $\omega(\alpha^2{}_i)$, $i \in I$ are fixed quantities and $\rho^s{}_i \in K^s$, the equality we get can be defined as

$$\sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^{(s)B} \omega(a_i^1) + \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^{(s)H} \omega(a_i^1) = \sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^{(s)B} \omega(a_i^2) + \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^{(s)H} \omega(a_i^2)\,. \tag{9}$$

Spreading (9) on the case of μ- equality of objects α^1 and α^2 we'll get finally

$$\sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^{(s)B} \omega(a_i^1) + \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^{(s)H} \omega(a_i^1) = \frac{\mu}{1-\mu} \sum_{\substack{i \in I_1 \\ \rho_i^s \in K^s}} \rho_i^{(s)B} \omega(a_i^2) + \frac{\mu}{1-\mu} \sum_{\substack{i \in I_2 \\ \rho_i^s \in K^s}} \rho_i^{(s)H} \omega(a_i^2)\,, \tag{10}$$

where $\rho_i^{(s)B}$, $\rho_i^{(s)H}$, $i \in I$ - correspondingly the upper and the lower borders of I-interval of weighting coefficients on s-step of algorithm. Equality (10) is an equivalent to equality (8).

So HWC on s+1 step will be equal

$$K^{s+1} = \prod_{i \in I_1} [\rho_i^{(s)H}, \rho_i^{(s+1)B}] \times \prod_{i \in I_2} \left[\rho^{(s+1)H}, \rho^{(s)B}\right]\,, \tag{11}$$

Equality (9) probes righteousness of statement 1: weighting coefficients defined with the help of the described above method limit HWC frames.

As only the fact of PMD preference is known, it is given in the form (3), so for defining vector component $\rho = (\rho_i,\ i \in I)$ we'll suggest the hypothesis about righteousness towards inequality of such a type:

$$\rho_{\substack{i \\ i \in I_1}} \omega(a^1{}_i) + \rho_{\substack{j \\ j \in I_2}} \omega(a^1{}_j) \leq \frac{\mu}{1-\mu} (\rho_{\substack{i \\ i \in I_1}} \omega(a^2{}_i) + \rho_{\substack{j \\ j \in I_2}} \omega(a^2{}_j))\,, \tag{12}$$

where $k = k_1 \cdot k_2$; k_1 – quantity of parameters with indexes i, $i \in I_1$; k_2 – quantity of parameters with indexes j, $j \in I_2$.

It is obvious that performing inequality (12) is a sufficient condition for performing inequality (7).

Let's pass over in equality system (12) to equalities and exclude k-n-1 equality according to the rule: in each step we exclude equality which add maximum to expression:

$$\max((\omega(a^1{}_i) - \omega(a^2{}_i),\ i \in I_1,\ \omega(a^2{}_j) - \omega(a^1{}_j),\ j \in I_2)\,.$$

The last condition means that such equalities are excluded which create unwarranted weights enlargement of one group of objects' parameters at the expense of the other group.

Let's add to the system n-1 of inequalities of type (12) as n-equality the condition of weights coefficients setting (5). Let's pass over to equalities and reorganize. We'll get finally the system of n-equality of the type:

$$\underset{i\in I_1}{\rho}\,{}_i(\omega(a^1{}_i)-\frac{\mu}{1-\mu}\omega(a^2{}_i))-\rho_j(\underset{j\in I_2}{\frac{\mu}{1-\mu}}\omega(a^2{}_j)-(\omega(a^1{}_i))=0\,, \tag{13}$$

$$\sum_{i\in I}\rho_i=1,\rho_i>0,i\in I\,.$$

From the equality system (13) we precisely define the components of weighting coefficients vector $\rho=(\rho_i,\ \ i\in I)$ which according to statement 1 limits in vector dimension R^n the intervals of weighting coefficients of objects' parameters.

Taking into consideration the information given above the correctness of the following statement is obvious.

Statement 2. The condition of objects selection ω^j, $j\in J$, from the set A^s is unbelongingness of HWC vector which passes through the coordinates beginning and point $\omega(\alpha^j)$, $\alpha^j\in A^s$, $j\in J$, namely $\rho(\omega(\alpha^j))\notin K^{(s+1)}$. Vector of weighting coefficients is defined according to the formula given in [Волкович, 1993]:

$$\rho=\rho(\omega(a^j))=\{\rho_i:\rho_i=\prod_{\substack{t\in I\\ t\neq i}}\omega(a^j{}_t)/\sum_{\substack{q\in I\\ l\neq q}}\prod_{l\in I}\omega(a^j{}_l)\}\,.$$

The person-computer procedure of localization of hyperparallelepiped of weighting coefficients is described in the following sequence.

Step 1. Pointing out the set of effective objects A^0 on the universal set A by one of the methods which are described in the article [Волкович, 1993]. The very first HWC is set as equal to single hypercube.

Step 2. PMD's choice of two objects α^1 and α^2 from the set of effective objects A^s in HWC K^s, s =1,2,... (step of HWC limiting) stating preference and equivalence.

Step 3. Constructing equation system of type (13). Finding solution of the equation system.

Step 4. Specifying HWC limits according to formula (11). If hypercube $K^{(s+1)}$ satisfies PMD's requirements it means the end of the procedure. Otherwise we pass over to the next step.

Step 5. Pointing out the set of effective objects $A^{(s+1)}$ ($A^{(s+1)}\subseteq A^{(s)}$) in HWC $K^{(s+1)}$ and presenting them to PMD for the choice of next two objects with stating for them the preference attitude.

Step 6. Uniting the objects, chosen by the expert on the previous step to the set of discussed objects and analysis on the given set of transitiveness. If the transitiveness is not destroyed then increase of iteration number s=s+1 and passing by to step 2. If the transitiveness is destroyed then the exclusion of these objects from the set of considered objects and passing by to step 6.

Conclusions

The suggested procedures do not demand the complete metrics of binary comparisons of objects and allow to restore the function of expert's value on the fuzzy binary relations' set. The reflection of weighted coefficients vector in the form of the intervals allows to present adequately the level of uncertainty in fuzzy models of decision making.

Bibliography

[Ларичев, 2002] Ларичев О.И. Свойства методов принятия решений в многокритериальных задачах индивидуального выбора // Автоматика и телемеханика, № 2, 2002. – С. 146-158.

[Борисов, 1989] Борисов А.Н., Алексеев А.В. и др. Обработка нечеткой информации в системах принятия решений. – М: Радио и связь, 1989. – 304 с.

[Пономаренко, 1994] Пономаренко О.І. Системні методи в економіці, менеджменті та бізнесі. – К.: Наукова думка, 1994. – 242 с.

[Волошин, 2003] Волошин А.Ф., Гнатиенко Г.Н., Дробот Е.В. Метод косвенного определения интервалов весовых коэффициентов параметров для метризированных отношений между объектами // Проблемы управления и информатики, 2003, № 2. – С. 83-89

[Волкович, 1993] Волкович В.Л., Волошин А.Ф., др. Модели и методы оптимизации надежности слождных систем. – К.: Наукова думка, 1993. – 312 с.

Author's Information

Elena Prisyazhnyuk – Kirovohrad Pedagogical University, Kirovohrad, Ukraine, e-mail: elena_drobot@ukr.net

THE NEW SOFTWARE PACKAGE FOR DYNAMIC HIERARCHICAL CLUSTERING FOR CIRCLES TYPES OF SHAPES

Tetyana Shatovska, Tetiana Safonova, Iurii Tarasov

Abstract: In data mining, efforts have focused on finding methods for efficient and effective cluster analysis in large databases. Active themes of research focus on the scalability of clustering methods, the effectiveness of methods for clustering complex shapes and types of data, high-dimensional clustering techniques, and methods for clustering mixed numerical and categorical data in large databases. One of the most accuracy approach based on dynamic modeling of cluster similarity is called Chameleon. In this paper we present a modified hierarchical clustering algorithm that used the main idea of Chameleon and the effectiveness of suggested approach will be demonstrated by the experimental results.

Keywords: Chameleon, clustering, hypergraph partitioning, coarsening hypergraph.

ACM Classification Keywords F.2.1 Numerical Algorithms and Problems

Introduction

The process of grouping a set of physical or abstract objects into classes of similar objects is called clustering. A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. A cluster of data objects can be treated collectively as one group in many applications. Data clustering is under vigorous development. Contributing areas of research include data mining, statistics, machine learning, spatial database technology, biology, and marketing. Owing to the huge amounts of data collected in databases, cluster analysis has recently become a highly active topic in data mining research. As a branch of statistics, cluster analysis has been studied extensively for many years, focusing mainly on distance-based cluster analysis. Active themes of research focus on the scalability of clustering methods, the effectiveness of methods for clustering complex shapes and types of data. Chameleon is a clustering algorithm that explores dynamic modeling in hierarchical clustering. In its clustering process, two clusters are merged if the interconnectivity and closeness between two clusters are highly related to the internal interconnectivity and closeness of objects within the clusters. The merge process based on the dynamic model facilitates the discovery of natural and homogeneous clusters and applies to all types of data as long as a similarity function is specified. Chameleon is derived based on the observation of the weakness of two hierarchical clustering algorithms: CURE and ROCK. CURE and related schemes ignore information about the aggregate interconnectivity of objects in two different clusters, whereas ROCK and related schemes ignore information about the closeness of two clusters while emphasizing their interconnectivity. In this paper, we present our experiments with hierarchical clustering algorithm CHAMELEON for circles cluster shapes with different densities using hMETIS program that used multilevel k-way partitioning for hypergraphs and a Clustering Toolkit package that merges clusters based on a dynamic model. In CHAMELEON two clusters are merged only if the inter-connectivity and closeness between two clusters are comparable to the internal inter-connectivity of the clusters and closeness of items within the clusters. The methodology of dynamic modeling of clusters is applicable to all types of data as long as a similarity